ABSTRACT OF THE DISCLOSURE

The invention provides a method for identifying a polypeptide that binds a ligand. The method includes the steps of (a) comparing a sequence of a polypeptide to

5 a sequence model for polypeptides that bind a ligand, wherein the sequence model comprises representations of amino acids consisting of a subset of amino acids, the subset of amino acids having one or more atom within a selected distance from a bound ligand in the polypeptides

10 that bind the ligand; and (b) determining a relationship between the sequence and the sequence model, wherein a correspondence between the sequence and the sequence model identifies the polypeptide as a polypeptide that binds the ligand.